09/673448
PCT/AU99/00306

5

Patent Office
Canberra

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PP 3129 for a patent by COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION filed on 23 April 1998.

WITNESS my hand this Fifth day of May 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

AUSTRALIA

Patents Act 1990

COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION

PROVISIONAL SPECIFICATION

Invention Title:

Diagnostic Assay

The invention is described in the following statement:

Field of the Invention:

This invention relates to an assay for diagnosis or prognosis of a disease or condition characterised by abnormal methylation of cytosine at a site or sites within the glutathione-S-transferase (GST) Pi gene and/or regulatory flanking sequences. In one particular application, the assay provides for the diagnosis or prognosis of prostate cancer.

Background of the Invention:

DNA METHYLATION IN MAMMALIAN GENOMES

The only established post-synthetic modification of DNA in higher animal and plant genomes is methylation of the 5' position of cytosine. The proportion of cytosines which are methylated can vary from a few percent in some animal genomes (1) to 30% in some plant genomes (2). Much of this methylation is found at CpG sites where the symmetrically positioned cytosines on each strand are methylated. In plant genomes, similar symmetrical methylation of cytosines at CpNpG (where N can be any base) is also common (3). Such sites of methylation have also been identified at low frequency in mammalian DNA (4).

Methylation patterns are heritable as the methylase enzyme recognises as a substrate, sites where a CpG dinucleotide is methylated on one strand but the corresponding C on the other strand is unmethylated, and proceeds to methylate it (5, 6). Fully unmethylated sites do not normally act as substrates for the enzyme and hence remain unmethylated through successive cell divisions. Thus, in the absence of errors or specific intervening events, the methylase enzyme enables the stable heritability of methylation patterns.

Extensive studies of gene expression in vertebrates have shown a strong correlation between methylation of regulatory regions of genes and their lack of expression (7). Most of such studies have examined only a limited number of restriction enzyme sites using enzymes which fail to cut if

their target sites are methylated. A far more limited number have been examined at all cytosine bases using genomic sequencing methods (8, 9).

BISULPHITE CONVERSION OF DNA

Treatment of single-stranded DNA with high concentrations of bisulphite followed by alkali leads to the selective de-amination of cytosine, converting it to uracil (10, 11). By contrast, 5-methyl cytosines (5meC) are resistant to this chemical de-amination. When bisulphite-treated DNA is copied by DNA polymerases, the uracils are read as if they were thymines and an adenine nucleotide incorporated, while 5meC is still read as a cytosine (a G being incorporated opposite). Thus, after a region of sequence is amplified by polymerase chain reaction (PCR), cytosines in the sequence which were methylated in the original DNA will be read as cytosines while unmethylated cytosines will be read as thymines (12, 13).

PCR AMPLIFICATION OF METHYLATED AND UNMETHYLATED DNA

In order to amplify bisulphite-treated DNA, primers are designed to anneal to the sequence produced after bisulphite treatment of the DNA. Since cytosines are converted to uracils, the base in the annealing primer will be an adenine rather than a guanine for the non-converted cytosine. Similarly, for the other primer of the pair, thymines replace cytosines. To permit quantification of levels of methylation in the target DNA, primers are normally chosen to avoid sites which may or may not be methylated (particularly CpG sites) and so may contain either a 5meC or a uracil after bisulphite treatment. Use of such non-selective primers allows both methylated and unmethylated DNAs to be amplified by PCR, providing for quantification of the level of methylation in the starting DNA population. The PCR-amplified DNA can be cut with an informative restriction enzyme, can be sequenced directly to provide an average measure of the proportion of methylation at any position or molecules may be cloned and sequenced (each clone will be derived from amplification of an individual strand in the initial DNA). Such studies have indicated that, while a population of molecules

may conform to an overall pattern of methylation, not all molecules will be identical and methylation may be found on only a fraction of molecules at some sites (reference 13,16).

SELECTIVE AMPLIFICATION OF METHYLATED DNA

Recently Herman *et al.* (14) described a variation of the bisulphite sequencing procedure to make it selective for the amplification of only methylated DNA. In this work, they reported the use of PCR primers which were designed to discriminate between the sequences produced after bisulphite-treatment of methylated and non-methylated target DNAs. Thus cytosines which formed part of a CpG site would not be bisulphite converted and would remain as cytosines in the methylated DNA but would be converted to uracils in the unmethylated target DNA. Primers utilising these differences were designed and used for the amplification of methylated DNA sequences from four tumour suppressor genes, p16, p15, E-cadherin and von Hippel-Lindau.

METHYLATION OF THE GLUTATHIONE-S-TRANSFERASE Pi GENE IN PROSTATE CANCER

Lee *et al.* (15) (US Patent No 5,552,277 and International Patent Application No PCT/US95/09050) demonstrated that expression of the glutathione-S-transferase (GST) Pi gene is lost in nearly all cases of prostate cancer. They further showed that in twenty cases examined, using Southern blotting, that this loss of expression was accompanied by methylation at a specific restriction enzyme site (*Bss*HII) in the promoter region of the gene. This methylation was not seen in normal prostate tissue or in a number of other normal tissues examined. In examining a prostate cancer cell line in which the GST-Pi gene is inactive, they also identified methylation at two other restriction enzyme sites, *Not*I and *Sac*II in the promoter region of the gene. Digestion of cell line DNAs with the enzymes *Msp*I and *Hpa*II, indicated that the correlation of DNA methylation with lack of expression was not maintained for these sites which were largely located downstream of

the transcription start site. The nature of the data makes it difficult to reach conclusions on the methylation status of individual *MspI/HpaII* sites. However, Lee *et al.* (18) were able to show that following *HpaII* digestion (which will cut at all unmethylated *HpaII* sites), a region of DNA containing twelve *HpaII* recognition sites could be amplified by PCR from tumour DNA, but not from normal prostate or leukocyte DNA. This indicates that some DNA molecules in prostate cancer are methylated at all these *HpaII* sites, while DNAs from normal prostate and leukocyte DNA must contain at least one of these sites unmethylated (as a single cut will render the region incapable of being amplified by PCR).

The present inventors have identified and developed an alternative method for detecting sites of methylation present in DNA from prostate cancer tissue but not present in DNA from normal tissue. The method relies on selective PCR amplification but does not require prior restriction with an informative restriction enzyme.

Disclosure of the Invention:

Thus, in a first aspect, the present invention provides a diagnostic or prognostic assay for a disease or condition in a subject, said disease or condition characterised by abnormal methylation of cytosine at a site or sites within the glutathione-S-transferase (GST) Pi gene and/or regulatory flanking sequences, wherein said assay comprises;

(i) isolating DNA from said subject,

(ii) exposing said isolated DNA to reactants and conditions for the amplification of a target region of the GST-Pi gene and/or regulatory flanking sequences which includes a site or sites at which abnormal cytosine methylation characteristic of the disease or condition occurs, the amplification being selective in that it only amplifies the target region if the said site or sites at which abnormal cytosine methylation occurs is/are indeed methylated, and

(iii) determining the presence of amplified DNA.

Since the amplification is designed to only amplify the target region if the said site or sites at which abnormal cytosine methylation (i.e. as compared to the corresponding site or sites of DNA from subjects without the disease or condition being assayed) occurs is/are methylated, the presence of amplified DNA will be indicative of the disease or condition in the subject from which the isolated DNA has been obtained. The assay thereby provides a means for diagnosing or prognosing the disease or condition in a subject.

Preferably, the disease or condition to be assayed is selected from cancers, especially hormone dependent cancers such as prostate cancer, breast cancer, cervical cancer and ovarian cancer or liver cancer. Most preferably, the disease or condition to be assayed is prostate cancer.

The step of isolating DNA may be conducted in accordance with standard protocols. The DNA may be isolated from any suitable body sample, such as cells from tissue, blood, semen or urine.

Prior to the amplifying step, the isolated DNA is preferably treated such that unmethylated cytosines are converted to uracil or another nucleotide capable of forming a base pair with adenine while methylated cytosines are unchanged or are converted to a nucleotide capable of forming a base pair with guanine. This treatment permits the design of primers which when used in polymerase chain reaction (PCR) amplification, will selectively amplify the target region if the said site or sites at which abnormal cytosine methylation occurs is/are methylated.

Preferably, following treatment and PCR amplification of the isolated DNA, a test is performed to verify that unmethylated cytosines have been efficiently converted to uracil or another nucleotide capable of forming a base pair with adenine, and that methylated cystosines have remained unchanged or efficiently converted to another nucleotide capable of forming a base pair with guanine.

Preferably, the treatment of the isolated DNA involves reacting the isolated DNA with bisulphite in accordance with standard protocols. As will be clear from the above discussion of bisulphite treatment, unmethylated cytosines will be converted to uracil whereas methylated cytosines will be unchanged. Verification that unmethylated cytosines have been converted to uracil and that methylated cystosines have remained unchanged may be achieved by;

(i) restricting an aliquot of the treated and amplified DNA with a suitable restriction enzyme(s) which recognise a restriction site(s) generated by or resistant to the bisulphite treatment, and

(ii) assessing the restriction fragment pattern by electrophoresis.

Alternatively, the verification may be achieved by differential hybridisation using specific oligonucleotides targeted to regions of the treated DNA where unmethylated cytosines would have been converted to uracil and methylated cytosines would have remained unchanged.

The amplifying step may be conducted in accordance with standard protocols for PCR amplification. Typically, the reactants will be suitable primers, dNTPs and thermostable DNA polymerase, and the conditions will be cycles of varying temperatures and durations to effect alternating denaturation of strand duplexes, annealing of primers and subsequent DNA synthesis.

As indicated above, selective PCR amplification with bisulphite-treated DNA may be achieved by using primers and conditions (e.g. high stringency conditions) that discriminate between a target region where the said site or sites at which abnormal cytosine methylation occurs is/are methylated and a target region where the said site or sites at which abnormal cytosine methylation occurs is/are unmethylated. Thus, for amplification only of a target region where the said site or sites at which abnormal cytosine methylation occurs is/are methylated, the primers used to anneal to the bisulphite-treated DNA (i.e. reverse primers) will include a guanine nucleotide(s) at a site(s) at which it will form a base pair with a methylated cytosine(s). Such primers will form a mismatch if the target region in the isolated DNA has unmethylated cytosine nucleotide(s) (which would have been converted to uracil by the bisulphite treatment) at the site or sites at which abnormal cytosine methylation occurs. The primers used for annealing to the opposite strand (i.e. the forward primers) will include a cytosine nucleotide(s) at any site(s) corresponding to site(s) of methylated cytosine in the bisulphite-treated DNA.

Preferably, the primers used for the PCR amplification are of 12 to 30 nucleotides in length and are designed to anneal to a sequence within the target region that includes two to four cytosine nucleotides that are abnormally methylated in the DNA of a subject with the disease or condition being assayed. In addition, the primers preferably include a terminal nucleotide that will form a base pair with a cytosine nucleotide (reverse primer), or the guanine opposite (forward primer), that is abnormally

methylated in the DNA of a subject with the disease or condition being assayed.

The step of amplifying is used to amplify a target region within the GST-Pi gene and regulatory flanking sequences. The regulatory flanking sequences may be regarded as the sequences 5' and 3' of the GST-Pi gene which include the elements that regulate, either alone or in combination with another like element, expression of the GST-Pi gene. Preferably, the regulatory flanking sequences consist of the 400 nucleotide sequence immediately 5' of the transcription start site and the 100 nucleotide sequence immediately 3' of the transcription stop site.

More preferably, the step of amplifying is used to amplify a target region within the region of the GST-Pi gene and regulatory flanking sequences defined by (and inclusive of) CpG sites -45 to +55 (wherein the numbering of the CpG sites is relative to the transcription start site). The numbering and position of CpG sites is shown in Figure 1.

For the diagnosis or prognosis of prostate cancer, the step of amplifying preferably amplifies a target region within the region of the GST-Pi gene and flanking regulatory sequences defined by (and inclusive of) nucleotides -43 to +53, more preferably, -43 to +10. However, within these target regions it is believed that there are CpG sites which show variability in methylation status in prostate cancer or are methylated in other tissues. Thus, in the case of the target region defined by (and inclusive of) CpG sites -43 to +10, it is preferred that the primers used for amplification be designed so as to minimise (i.e. by use of redundant primers or by avoidance of the sites) the influence of CpG sites -36, -32, -23, -20, -14 and a polymorphic region covering site -33.

For discriminating between early (unmethylated) prostate cancer and later (methylated) prostate cancer, the step of amplifying preferably amplifies a target region within the region of the GST-Pi gene and flanking regulatory sequences defined by (and inclusive of) CpG sites +13 to +53.

For assays using body samples other than cells from prostate tissue, e.g. blood, it is preferred to restrict the amplification so as to amplify a target region that does not include the region of the GST-Pi gene and flanking regulatory sequences defined by (and inclusive of) CpG sites -7 to +7, since this may lead to false positives (e.g. where liver disease cells are present in the samples).

Most preferably, the step of amplifying involves PCR amplification using pairs of forward and reverse primers selected from the following groups:

Forward Primers (i.e. anneal to the 5' end of the target region)

CGCGAGGTTTTCGTTGGAGTTTCGTCGTC

CGTTATTAGTGAGTACGCGCGGTTC

YGGTTTTAGGGAATTTTTTTTCGC

YGGYGYGTTAGTTYGTTGYGTATATTTC

GGGAATTTTTTTTCGCGATGTTTYGGCGC

TTTTTAGGGGGTTYGGAGCGTTTC

GGTAGGTTGYGTTTATCGC

Reverse Primers (i.e. anneal to the extension of the forward primer)

TCCCATCCCTCCCCGAAACGCTCCG

GAAACGCTCCGAACCCCCTAAAAACCGCTAACG

CRCCCTAAAATCCCCRAAATCRCCGCG

ACCCCRACRACCRCTACACCCCRAACGTCG

CTCTTCTAAAAAATCCCRCRAACTCCCGCCG

AAAACRCCCTAAAATCCCCGAAATCGCCG

AACTCCCRCCGACCCCAACCCCGACGACCG

AAAAATTCRAATCTCTCCGAATAAACG

AAAAACCRAAATAAAAACCACACGACG

where Y is a mixture of C and T, and R is a mixture of A and G.

Such primer pairs can be used to selectively amplify only a target region where the said site or sites at which abnormal cytosine methylation occurs is/are methylated.

In a second aspect, the present invention provides primers (sequence shown in the 5' to 3' direction) comprising a nucleotide sequence selected from the group consisting of:

CGCGAGGTTTTCGTTGGAGTTTCGTCGTC

CGTTATTAGTGAGTACGCGCGGTTC

YGGTTTTAGGGAATTTTTTTTCGC

YGGYGYGTTAGTTYGTTGYGTATATTTC

GGGAATTTTTTTTCGCGATGTTTYGGCGC

TTTTTAGGGGGTTYGGAGCGTTTC

GGTAGGTTGYGTTTATCGC

AAAAATTCRAATCTCTCCGAATAAACG

AAAAACCRAAATAAAAACCACACGACG

TCCCATCCCTCCCCGAAACGCTCCG

GAAACGCTCCGAACCCCCTAAAAACCGCTAACG

CRCCCTAAAATCCCCRAAATCRCCGCG

ACCCCRACRACCRCTACACCCCRAACGTCG

CTCTTCTAAAAAATCCCRCRAACTCCCGCCG

AAAACRCCCTAAAATCCCCGAAATCGCCG

AACTCCCRCCGACCCCAACCCCGACGACCG,

where Y is a mixture of C and T, and R is a mixture of A and G.

The identification of specific sites of abnormal methylation of cytosines within the GST-Pi gene and regulatory flanking sequences from prostate cancer cells, enables the design and production of specific agents such as oligonucleotides which may promote the demethylation of one or more of the abnormally methylated cytosines. Such agents may thereby bring about the reactivation of GST-Pi gene expression and may therefore be of significant therapeutic value. Similar agents may also be designed and

produced for the therapeutic treatment of other diseases and conditions characterised by abnormal cytosine(s) methylation within the GST-Pi gene and/or regulatory flanking sequences.

Thus, in a third aspect, the present invention provides an agent for the treatment of a disease or condition characterised by abnormally methylated cytosine(s) within the GST-Pi gene and/or regulatory flanking sequences, wherein said agent promotes the specific demethylation of said abnormally methylated cytosine(s).

Preferably, the agent is for the treatment of a disease or condition selected from cancers, especially hormone-dependent cancers such as prostate cancer, breast cancer, cervical cancer and ovarian cancer or liver cancer. Most preferably, the agent is for the treatment of prostate cancer.

Said agent may promote the specific demethylation of some or all of the abnormally methylated cytosine(s). For prostate cancer, the agent preferably promotes demethylation of some of all of the abnormally methylated cytosines within the region of the GST-Pi gene and regulatory flanking sequences defined by (and inclusive of) CpG sites -43 to +53, more preferably, -43 to +10 (wherein the numbering of the CpG sites is relative to the transcription start site).

The terms "comprise", "comprises" and "comprising" as used throughout the specification are intended to refer to the inclusion of a stated component, feature or step or group of components, features or steps with or without the inclusion of a further component, feature or step or group of components, features or steps.

The invention will now be further described with reference to the accompanying figures and following, non-limiting example.

Example: Diagnostic assay for prostate cancer involving detection of methylated GST-Pi DNA

METHODS

(1) Treatment of DNA with bisulphite

DNA for assaying was isolated from suitable sources by standard protocols and treated with bisulphite by well known methods (12, 13, 16).

(2) Characterisation of Methylation of Individual Sites in DNA

In order to determine the methylation status of individual cytosine nucleotides in target and non-target DNAs and to identify differences between them, bisulphite-modified DNA was amplified by PCR using primers designed to minimise the possibility that the methylation status of a particular CpG site will influence primer annealing and subsequent amplification (12, 13, 16).

(3) Design of Selective Primers

Based on the sequencing information, primers for use in the assay were designed to maximise the possibility that the methylation status of a particular CpG site would influence primer annealing and subsequent amplification. Specifically, the design principles followed (described for the "forward" PCR primer where the primer contains the same C to T (or U) conversions as would occur in the bisulphite-treated DNA), are listed below at (a) to (d). For the "reverse" primer, which anneals to the converted strand, A's replace G's at positions opposite converted C's

(a) That primers should cover sequence regions which contain a number of C's. Conversion of unmethylated C's to U's provides for discrimination between molecules which have undergone efficient bisulphite conversion and molecules in which C's have not reacted (e.g. because not completely dissolved or containing regions of secondary structure).

(b) That at least one, but preferably two to four, of the C's in the regions should be C's (generally at CpG sites) known to be methylated in a high proportion of the DNA to be detected (i.e. target DNA). Thus, these C's will

remain C's in the target DNA while being converted to U's in the non-target DNA. A primer which is designed to be the exactly equivalent of the bisulphite-converted methylated DNA will contain a mismatch at each of the positions of an unmethylated C which has been converted to a U in an unmethylated DNA. The more mismatches that are present, the greater the differential hybridisation stability of the primers will be and hence the greater the selective difference in PCR.

(c) That the 3' terminal base of the primer should preferably be a C corresponding to a C known to be methylated in the target DNA (normally part of a CpG dinucleotide). Correct pairing with the terminal base of the primer will provide for highly selective priming of target sequences compared with unmethylated background sequences which will form a C:A mismatch.

(d) That at positions where it is known that methylation occurs in only a fraction of molecules in the methylated target DNA or where it is known to vary between target DNAs (e.g.. in different tumour samples), redundancy can be incorporated into the primers to allow for amplification of either C or T from the target DNA. This same approach can be used if polymorphisms are known to exist in the primer region.

(4) Verification of Selective Target Sequence Amplification

An important principle of the method utilised in the assay is that the amplified PCR band is analysed to verify that it has been derived from DNA which has been fully bisulphite-converted (i.e. C's not methylated in the original DNA have been converted to U's and amplified as T's) and to further confirm that the amplified DNA is confirmed as having been derived from the specific target DNA sequence and has the expected methylation profile (i.e. 5meC's not converted to T's). Methods for determining this include:

(a) Using restriction enzyme digestion.

In order to verify complete conversion, particular restriction enzymes can be used to cut the DNA. The sequence recognition sites should have the

property that they contain no C's and are present in the sequence of the amplified strand after but not before bisulphite treatment. Thus, the conversion of one or preferably two or more C's to U's and their amplification as T's in the PCR product should produce a new restriction site. Useful enzymes are shown in italics in Table 1 below.

In order to verify that the target DNA sequence amplified was specifically methylated, use can be made of restriction enzyme sites whose only C nucleotides are found as CpG dinucleotides and which, if the sequence was methylated, would remain as CpG's in the PCR products. Examples of such enzymes are shown in bold in Table 1 below. *Bsm*BI, which cuts the non-symmetrical sequence GAGACG can also be used.

In some instances, enzymes which contain a C as an outer base in their recognition sequence can be used for verification of methylation: e.g. *Eco*RI (GAATTC) for a GAATTCG sequence or *Sau*3AI (GATC) for a GATCG sequence (bold and underlined in Table 1). If a site such as one of the above is present in the predicted methylated, fully bisulphite-converted DNA then the enzyme will cut the DNA only if the original CpG dinucleotide was methylated, confirming the amplification of a methylated region of DNA. Some of the enzymes (bold and underlined in Table 1) have the potential to be used both for monitoring efficient conversion and CpG methylation.

(b) Differential hybridisation to specific oligonucleotides.

Differential hybridisation to specific oligonucleotides can be used to discriminate that the amplified DNA is fully reacted with bisulphite and of the expected methylation profile. To demonstrate complete conversion, a pair of oligonucleotides corresponding to the same region within the amplified sequence is prepared. One oligonucleotide contains T's at all C's which should be converted by bisulphite, while the other contains C's in these positions. The oligonucleotides should contain two or three of such discriminatory C's and conditions be determined which provide for selective hybridisation of each to its target sequence. Similar oligonucleotides with C

or T at CpG sites and T's replacing all non-CpG C's are used to determine whether the specific CpG sites are methylated. The oligonucleotides can be used for direct hybridisation detection of amplified sequences or used to select out target molecules from the PCR-amplified DNA population for other detection methods. An array of such oligonucleotides on a DNA sequencing chip can be used to establish the sequence of the amplified DNA throughout the sequence region.

(c) Single nucleotide primer extension (SNuPE).

The technique of single nucleotide primer extension can be applied to the PCR products to determine whether specific sites within the amplified sequence contain C or T bases. In this method, a primer abutting the position of interest is annealed to the PCR product and primer extension reactions performed using either just dCTP or just dTTP. The products can be separated by gel electrophoresis and quantitated to determine the proportion of each nucleotide in the population at that position. Primers should be designed to quantitate conversion of C's in CpG sites and control C's which should not be methylated. More than one primer can be included in a single reaction and/or run in the same gel track as long as their sizes can be clearly distinguished.

(d) Fluorescent Real-time Monitoring of PCR.

Oligonucleotides internal to the amplified region can be used to monitor and quantify the amplification reaction at the same time as demonstrating amplification of the correct sequence. In the Fluorogenic 5' Nuclease PCR assay (19) the amplification reaction is monitored using a primer which binds internally within the amplified sequence and which contains both a fluorogenic reporter and a quencher. When this probe is bound to its target DNA it can be cleaved by the 5' nuclease activity of the Taq polymerase, separating the reporter and the quencher. By utilising in the assay an oligonucleotide which is selective for the fully bisulphite-converted sequence (and/or its methylation state) both the level of amplification and its specificity can be monitored in a single reaction.

METHYLATION SEQUENCE PROFILE OF TARGET AND NON-TARGET GST-Pi DNA

Figure 1 shows the organisation of the GST-Pi gene and the regions for which genomic sequencing was used to determine the methylation status of DNA isolated from prostate cancer tissue or cell lines and from normal prostate or other tissues. Also shown, within the boxes is the sequence of each amplified region, with all the CpG sites indicated and numbered relative to the position of the transcription start site. Sequence analysis demonstrated that there was an additional CpG dinucleotide (+9) not predicted from the published sequence. Also identified in the regions sequenced was a polymorphism which is present in a significant fraction of the samples studied. The polymorphic allele does not contain CpG site -33 (see Figure 3). Table 2 lists the sequences and positions of the non-selective primers used for amplification (Table 2-1) and direct sequencing (Table 2-2) of bisulphite-converted DNA. DNA isolated from normal prostate tissue, prostate cancer tissue, prostate cancer-derived cell lines and other tissues was bisulphite treated and PCR reactions done by standard procedures (13). PCR products were either digested with informative restriction enzymes, sequenced directly (17), or individual molecules cloned and sequenced by standard procedures.

Figure 2 shows the methylation status at each CpG site determined in the DNAs; the level of methylation detected at each site is shown, none (-), up to 25% (+), 26-50% (++), 51-75% (+++) and 76-100% (++++). The Gleason Grade of tumour samples is also shown.

In Figure 2a, the methylation status of sites in prostate cancer cell lines, prostate cancer tissue samples and matched normal prostate tissue are shown. for the core promoter regions through to the 3' end of the gene (covering sites -28 to 103). It can be seen that in normal prostate tissue, the core promoter region is unmethylated at all sites and that this lack of methylation extends through the region flanking the promoter to site +33.

Results of restriction enzyme digests of bisulphite-converted, PCR-amplified DNA indicate that this lack of methylation includes sites +52 and +53. However, in the regions further downstream which were analysed, sites +68 to +74 and +96 to +103, DNA from normal prostate tissue was heavily methylated. Analysis of the prostate cancer cell line LNCaP and prostate cancer tissue samples demonstrates extensive methylation of the core promoter region; variations in the overall level of methylation probably reflect the presence of different levels of normal cells within the tumour samples. DNA from one cancer sample (2AC) was found to be completely unmethylated and in contrast to the other tumour samples this tumour was found by immunohistochemistry to still be expressing GST-Pi. Sequencing of the region flanking the core promoter in the LNCaP cell line and tumour DNAs BC and CC showed that methylation extended through to site +33 and further restriction enzyme analysis showed that methylation included sites +52 and +53. For one tumour sample, DC, methylation did not extend beyond the core promoter region and sites +13 to +33, as well as sites +52 and +53 were found to be unmethylated. It is notable that this tumour was of Gleason Grade 2+2, the lowest grade tumour among those analysed. For all tumour DNA samples, as for the normal DNA, the downstream region of the gene, sites 68 to 74 and 96 to 103 , were heavily methylated. Within the promoter regions which were methylated in the cancer, but not normal, tissue specific individual sites were evident which were either unmethylated or methylated to a much lower degree than surrounding methylated sites. These include sites -22 and -23 (XC), -20 (PC3 lines, XC &WC), -14 (PC3, XC &WC), +24 (PC3-M & MM2, CC), +25 (LNCaP, PC3-MM2, CC).

Figure 2b compares the methylation state of the core promoter region and sequences upstream of it in normal prostate tissue DNA and in DNA isolated from a number of other normal tissues. Sequences from the PCR fragment upstream of the core promoter were determined by cloning and sequencing as the region is refractory to direct sequencing. For the cancer

samples, the level of methylation shown is as a proportion of those clones which were methylated (about 50% of the total clones in both cases). In normal prostate as well as in all other tissues there is extensive methylation of CpG sites upstream of the AT-rich repeat. Downstream of the repeat (from CpG site -43) minimal methylation was seen in all tissues except in liver, where there was significant methylation of sites -7 through to +7. Sequences upstream of the core promoter were found to be heavily methylated in the prostate cancer DNA analysed, though again specific sites were undermethylated; site -32 in cancers B and D and site -36 in cancer B.

The data thus identify a region of DNA, stretching from 3' of the polymorphic repeat region, (CpG site -43) to sites +52 and +53, which is not methylated in normal prostate tissue but is normally highly methylated in prostate cancer. In one cancer sample (D, the cancer of lowest Gleason Grade) the region from CpG sites +13 to +53 was not methylated. The more restricted region extending from CpG site -43 to CpG site +10 was methylated in all of the prostate cancer DNAs which showed promoter methylation. Methylation of part of the promoter region (CpG sites -7 to +7) was also seen in one normal tissue (liver) examined.

This information is critical in defining regions of the GST-Pi gene which can be used for the development of methods and reagents for the selective detection of prostate cancer cells. Thus the region from CpG sites -43 to +53 lying within the boundary of regions methylated in normal prostate tissue can be used for the design of primers to detect cancer-specific methylation in prostate tissue. The region from site -43 to +10 is preferred for the detection of a higher proportion of cancers. The region from sites +13 to +53 may be used to distinguish early (unmethylated) cancer from later (methylated cancer). For assay of other samples, such as blood, for the presence of prostate cancer cells it is preferred to restrict the region chosen to exclude sites -7 to +7; for example in the case of liver disease liver cells may be present in the blood which could lead to a false positive for the presence

of prostate cancer cells. The information is used in the design of methylation selective primers for detection of prostate cancer DNA below. The methylation profile of the GST-Pi gene may also be used for development of strategies for re-activation or prevention of inactivation of GST-Pi expression in therapeutic or preventative strategies for prostate cancer.

DESIGN AND USE OF SELECTIVE PRIMERS FOR DETECTION OF METHYLATED GST-Pi DNA

Sequence primers for the detection of methylated GST-Pi sequences from three regions, upstream of the core promoter (primers CGPS-5,6,7,8,9,11,12 &13), partially encompassing the core promoter (primers CGPS-1 to 4) and further downstream from the core promoter (primers CGPS-21 to 24) are shown in Table 3 below. The sequence and derivation of primers for the upstream region are shown for example in Figure 3 (from CpG site -43 to CpG site +10). The common polymorphism encompassing CpG site -33 is shown above the sequence (p). Underneath is shown the sequence of the derived strand after conversion of cytosines to uracil. The derived strand is shown either assuming all CpGs are methylated (B-M) or that none are (B-U). Below this are specific primers designed to selectively amplify the methylated sequence. It can be seen that all primers are designed to match perfectly to the modified, methylated template, but contain mismatches to the template derived from unmethylated DNA or the original unmodified DNA. Primers CGPS-5, 8, 11, 12 & 13 are designed to avoid the polymorphic region and CpG sites which show a lower frequency of methylation in prostate cancer DNA. The underlined T's in the forward primers (and A's in the reverse primers) derive from bisulphite conversion of C's and provide discrimination against amplification of DNA which has not been efficiently converted by the bisulphite treatment. The bold C's in the forward primers (and G's in the reverse primers) are parts of CpG sites and will pair with DNA derived from methylated sequences but form mismatches to DNA derived from unmethylated sequences. Redundancy is included in some positions, Y

(= mix of C & T) in forward primers and R (= mix of A & G) in reverse primers to allow pairing independent of methylation status. This can allow for certain sites where the frequency of methylation within or between tumour samples is variable (eg. site -14). Forward and reverse primers for specific selective amplification of methylated GST-Pi sequences are shown in Table 3 below.

In Figure 4 the results of amplification of bisulphite modified DNAs from a variety of tissues using two sets of PCR primers are shown. Amplification reactions in Figure 4a, Panel A (region covering the transcription start site) used the CGPS-1 and 3 as outer primers and CGPS-2 and 4 as inner primers. In Figure 4a, Panel B and Figure 4b and 4c, the outer primer pair, CGPS-5 and CGPS-8 which encompass the region from CpG site -39 to -16 were used for first round amplification, followed by a second round of amplification with the CGPS-6 and CGPS-7 primers, amplifying a 90bp fragment covering CpG sites -36 to -23.

RESULTS

For the primers covering the core region, Figure 4a Panel A, the correct band (arrowed) is present in the positive control (cancer B) and in DNA from prostate samples of two men not diagnosed with prostate cancer. Bands are also seen in DNA from a bone marrow and blood sample as well as in liver DNA from subjects with no known prostate cancer.

For the upstream amplification, Figure 4a Panel B, it can be seen that a PCR product is not produced upon amplification of DNA from a range of healthy tissues nor from blood DNA samples of subjects without prostate cancer; a band is produced from the positive control, cancer B, DNA. DNA isolated from one normal prostate did not show an amplified fragment, while that from the prostate of an 82 year old did. It is possible that this person had undiagnosed prostate cancer. DNA isolated from five other samples of normal prostate tissue from patients without prostate cancer did not give rise to a PCR product (Figure 4c Panel B).

In Figure 4b, PCR reactions are shown for tissue samples from patients with prostate cancer: for each sample DNA was isolated from a region identified as containing cancer and from another region identified as grossly normal. In all cases, a clear PCR band is produced from amplification from cancer DNA. Two of these were cases where the proportion of methylated DNA was insufficient to be detected using primers designed to prime equivalently on methylated and unmethylated DNA. For grossly normal tissue, the PCR band is either absent or present in substantially lower amounts. The presence of a band in some "normal" samples could derive from a low level of cancer cells in the sample. Amplification of DNA from samples of blood obtained from the abdominal cavity during surgery showed that it was possible to detect methylated GST-Pi sequences in a number of them. Samples of peripheral blood isolated from three patients with known metastatic disease (Fig 4c, Panel B) demonstrated the presence of amplifiable, methylated GST-Pi sequences. PCR products were produced from amplification of DNA from the LNCaP and DU145 prostate cancer cell lines, but not from the PC-3 series of lines. This latter result could be because of the low level of methylation in the upstream promoter region in PC-3 cells, but a major contributing factor is likely to be a lack of priming by the CGPS-6 primer as PC-3 only contains the variant allele of the GST-Pi gene.

The results thus show that the specific primer set selectively amplifies DNA from prostate cancer cell lines, prostate cancer tissue, and from the blood of patients with known metastatic disease.

Methylated GST-Pi sequences are also detected in DNA isolated from some tumour-derived cell lines of non-prostatic origin: HeLa, a cervical carcinoma, and HepG2, a liver carcinoma (Figure 4c, Panel B).

CONFIRMATION OF CORRECT AMPLIFICATION

The specific oligonucleotides described below can be used to confirm that the amplified PCR fragment derives from DNA in which all non-methylated C's had been converted to U's. These for the upstream pcr region

can be used with PCR products from all combinations of the CGPS-5, 6, 11, 7, 8, 9, 12 and 13 forward and reverse primers. Those for the downstream region can be used with PCR products of the CGPS-21, 22, 23 and 24 primers. A biotinylated version of the conversion-specific olignucleotide can also be used for the selective and specific capture from solution of the PCR products generated using these primer pairs, or the appropriately labelled oligonucleotide can be use for real-time monitoring of specific PCR fragment amplification. PCR products from amplification of bisulphite-converted DNA routinely have one strand containing a very high proportion of T's and the other of A's. Because of this, it is possible to use oligo dT (or oligo dA) as a generic conversion specific oligonucleotide, the annealing conditions being varied to optimise discrimination of converted and non-converted DNA for each PCR fragment.

Upstream PCR region

Conversion specific oligonucleotide: 5'-AAACCTAAAAAATAAACAAACAA

Non-converted sequence oligonucleotide: 5'-GGGCCTAGGGAGTAAACAGACAG

Downstream PCR region:

Conversion specific oligonucleotide: 5'-CTTGGCATCCTCCCCCGGGCTCCAG

Non-converted sequence oligonucleotide: 5'-TTTGGTATTTTTTTTCGGGTTTTAG

References:

1. H. S. Shapiro, *Content of 6-methylaminopurine and 5-methylcytosine in DNA*. G. D. Fasman, Ed., Handbook of Biochemistry, Selected Data for Molecular Biology: Nucleic Acids. (CRC Press, Boca Raton FL, 1975).

2. R. L. Adams, Burdon, R. H., Molecular Biology of DNA methylation (Springer Verlag, New York, 1985).

3. Y. Gruenbaum, T. Naveh-Many, H. Cedar, A. Razin, *Nature* **292**, 860-862 (1981).

4. S. J. Clark, J. Harrison, M. Frommer, *Nat Genet* **10**, 20-27 (1995).

5. R. Holliday, J. E. Pugh, *Science* **187**, 226-232 (1975).

6. A. D. Riggs, *Cytogenet. Cell Genet.* **14**, 9-25 (1975).

7. M. Graessmann, A. Graessmann, in *DNA Methylation: Molecular Biology and Biological Significance* J. P. J. H. P. Saluz, Ed. (Birkhauser Verlag, Basel, Switzerland, 1993) pp. 404-424.

8. C. Stirzaker, et al., *Cancer Res* **57**, 2229-2237 (1997).

9. K. D. Tremblay, K. L. Duran, M. S. Bartolomei, *Mol Cell Biol Molecular and Cellular Biology* **17**, 4322-4329 (1997).

10. R. Y.-H. Wang, Gehrke, C.W. and Ehrlich, M., *Nucl. Acids Res.* **8**, 4777-4790 (1980).

11. R. Shapiro, B. Braverman, J. B. Louis & R. E. Servis, *J. Biol. Chem.* **248**, 4060-4064 (1973).

12. M. Frommer, McDonald, L.E., Millar, D.S., Collis, C..M., Watt, F., Grigg, G.W., Molloy, P.L. & Paul, C. L., *Proc. Natl. Acad. Sci. USA.* **89**, 1827-1831 (1992).

13. S. J. Clark, Harrison, J., Paul, C.L., & Frommer, M., *Nucl. Acids Res.* **22**, 2990-2997 (1994).

14. J. G. Herman, J. R. Graff, S. Myohanen, B. D. Nelkin &S. B. Baylin, *Proc Natl Acad Sci USA* **93**, 9821-9826 (1996).

15. W. H. Lee, et al., *Proc Natl Acad Sci USA* **91**, 11733-11737 (1994).

16. S. J. Clark, M. Frommer, in *Laboratory Methods for the detection of mutations and polymorphisms in DNA*. G. Taylor, Ed. (CRC press, NewYork, 1997) pp. 151-162.

17. C. L. Paul, S. J. Clark, *BioTechniques* **21**, 126-133 (1996).

18. W.H. Lee, W.B. Isaacs, G.S. Bova & W.G. Nelson, *Cancer Epidemiol Biomarkers Prev.* **6**, 443-450 (1997).

19. P.M. Holland, R.D. Abramson, R. Watson & D.H. Gelfland, *Proc. Natl. Acad. Sci. USA* **88**, 7276-7280 (1991).

TABLE 1

	AATT	TTAA	ATAT	ACGT	CGCG	GATC	TCGA	AGAG
↓ oooo	*Tsp509 I*			**Mae II**		**Sau3A**		
↓ oooo		*Mse I*					**Taq I**	
↓ oooo					**BstUI**			
↓ oooo				**Tai I**				
↓ AoooooT	*Apo I*							
↓ AoooooT		*Ase I*						
↓ AoooooT			*Ssp I*					
↓ TooooA							**BstBI**	
↓ TooooA				**SnaBI**	**Nru I**			
↓ CooooG						**Pvu I**		
↓ GooooC	**EcoRI** *Apo I*							

Table 2.1 Primers for PCR Amplification of the Bisulphite-Modified GST-Pi Gene

PCR #	Target	Primer Name	Primer Type	Primer 5' 3'	Target size (bp)	Anneal oC	Genomic Position
	Upstream	GST-1	Outer	TTATGTAATAAATTTGTATATTTTGTATATG	646	50/50	381-411
-1	Top Strand	GST-25	Inner	TGTAGATTATTTAAGGTTAGGAGTT	499	50/50	495-519
	DNA	GST-3	Inner	AAACCTAAAAAATAAACAAACAACAAA	499	50/50	967-993
		GST-4	Outer	AAAAAACCTTTCCCTCTTTCCCAAATCCC	646	50/50	999-1027
	Exon 1	GST-9	Outer	TTTGTTGTTTGTTTATTTTTTAGGTTT	346	45/50	967-993
1	Top Strand	GST-11	Inner	GGGATTTGGGAAAGAGGGAAAGGTTT	307	45/50	999-1025
	DNA	GST-12	Inner	ACTAAAAACTCTAAACCCCATCCC	307	45/50	1280-1303
		GST-10	Outer	AACCTAATACTACC TTAACCCCAT	346	45/50	1304-1329
	Exon 1	GST-B1	Outer	AATCCTCTTCCTACTATCTATTTACTCCCTAAA	387	50/55	958-990
2	Bottom	GST-B2	Inner	AAAACCTAAAAAAAAAAAAAAAACTTCCC	314	50/55	999-1027
	Strand	GST-B3	Inner	TTGGTTTTATGTTGGGAGTTTTGAGTTTT	314	50/55	1285-1313
	DNA	GST-B4	Outer	TTTTGTGGGGAGTTGGGGTTTGATGTTGT	387	50/55	1317-1345
	Exon 2/Exon 3	GST-13	Outer	GGTTTAGAGTTTTTAGTATGGGGTTAATT	691	45/50	1287-1315
3	Top Stand	GST-14	Inner	TAGTATTAGGTTAGGGTTTT	603	45/50	1318-1337
	DNA	GST-15	Inner	AACTCTAACCCTAATCTACCAACAACATA	603	45/50	1920-1892
		GST-16	Outer	CA AAAAACTTTAAATAAACCCTCCTACCA	691	45/50	1978-1950
	Exon 5	GST-30	Outer	GTTTTGTGGTTAGGTTGTTTTTTTAGGTGTTAG	340	55/60	2346-2376
4	Top Stand	GST-31	Inner	GTTTTGAGTATTTGTTGTGTGGTAGTTTTT	265	40/45	2381-2416
	DNA	GST-32	Inner	TTAATATAAATAAAAAAAATATATTTACAA	265	40/45	2617-2646
		GST-33	Outer	CAACCCCCAATACCCAACCCTAATACAAATACTC	340	55/60	2653-2686
	Exon 7	GST-26	Outer	GGTTTTAGTTTTTGGTTGTTTGGATG	347	50/55	3845-3869
5	Top Stand	GST-27	Inner	TTTTTTTGTTTTTAGTATATGTGGGG	287	50/55	3874-3899
	DNA	GST-28	Inner	ATACTAAAAAAACTATTTTCTAATCCTCTA	287	50/55	4161-4132
		GST-29	Outer	CCAAACTAAAAACTCCAAAAAACCACTAA	347	50/55	4192-4164

Bases arising due to C to U conversion by bisulphite treatment are shown in bold

Table 2.2 Primers for Direct Sequencing of Amplified GST-Pi Gene PCR Fragments

PCR #	Target	Primer Name	Primer Type	Primer 5' 3'	Target size (bp)	Anneal oC	Genomic Position
1	Exon 1 Top Strand DNA	GST-11	M13	TGTAAAACGACGGCCAGTGGGATTTGGGAAAGAGGGAA	307	45/50	1003-1026
		GST-12	Biotin	BioACTAAAAACTCTAAACCCCATCCC	307	45/50	1288-1313
2	Exon 1 Bottom Strand DNA	GST-B2	M13	TGTAAAACGACGGCCAGTTGTTGGGAGTTTTGAGTTTT	314	50/55	999-1027
		GST-B3	Biotin	BioAAAACCTAAAAAAAAAAAAAAAACTTCCC	314	50/55	1285-1313
3	Exon2/3 Top Strand DNA	GST-14	M13	TGTAAAACGACGGCCAGTTAGTATTAGGTTA	603	45/50	1317-1337
		GST-15	Biotin	BioAACTCTAACCCTAATCTACCAACAACATA	603	45/50	1920-1892
4	Exon4/5 Top strand DNA	GST-31	M13	TGTAAAACGACGGCCAGTGTTTTGAGTATTTGTTGTG	265	55/60	2381-2410
		GST-32	Biotin	BioTTAATATAAATAAAAAAAATATATTTTACAA	265	55/60	2617-2646
5	Exon 7 Top strand DNA	GST-27	M13	TGTAAAACGACGGCCAGTGTTTTTAGTATATGTGG	287	50/55	3874-4132
		GST-28	Biotin	BioATACTAAAAAAACTATTTTCTAATCCTCTA	287	50/55	4161-4164

Extensions on "M13" primers for annealing of sequencing primer is underlined.

TABLE 3

Primer	Forward or Reverse	Primer Sequence (5'-3')	Co-ordinates	CpG sites
CGPS-1	F	CGCGAGGTTTTCGTTGGAGTTTCGTCGTC	1210-1239	-3 to +3
CGPS-2	F	CGTTATTAGTGAGTACGCGCGGTTC	1247-1271	+4 to +8
CGPS-3	R	TCCCATCCCTCCCCGAAACGCTCCG	1428-1452	+21 to +23
CGPS-4	R	GAAACGCTCCGAACCCCCTAAAAACCGCTAACG	1406-1438	+19 to +23
CGPS-5	F	YGGTTTTAGGGAATTTTTTTTCGC	894-917	-39 to -37
CGPS-6	F	YGGYGYGTTAGTTYGTTGYGTATATTTC	925-952	-36 to -31
CGPS-11	F	GGGAATTTTTTTTCGCGATGTTTYGGCGC	902-930	-38 to -34
CGPS-7	R	CRCCCTAAAATCCCCRAAATCRCCGCG	1038-1064	-23 to -27
CGPS-8	R	ACCCCRACRACCRCTACACCCCRAACGTCG	1077-1106	-16 to -21
CGPS-9	R	CTCTTCTAAAAAATCCCRCRAACTCCCGCCG	1113-1143	-12 to -15
CGPS-12	R	AAAACRCCCTAAAATCCCCGAAATCGCCG	1040-1068	-23 to -26
CGPS-13	R	AACTCCCRCCGACCCCAACCCCGACGACCG	1094-1123	-14 to -18
CGPS-21	F	TTTTTAGGGGGTTYGGAGCGTTTC	1415-1438	+21 to +23
CGPS-22	F	GGTAGGTTGYGTTTATCGC	1473-1491	+26 to +28
CGPS-23	R	AAAAATTCRAATCTCTCCGAATAAACG	1640-1666	+36 to +34
CGPS-24	R	AAAAACCRAAATAAAAACCACACGACG	1676-1703	+39 to +37

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

Dated this day of 1998.

COMMONWEALTH SCIENTIFIC
AND INDUSTRIAL RESEARCH
ORGANISATION
Patent Attorneys for the Applicant:

F B RICE & CO

FIG. 1

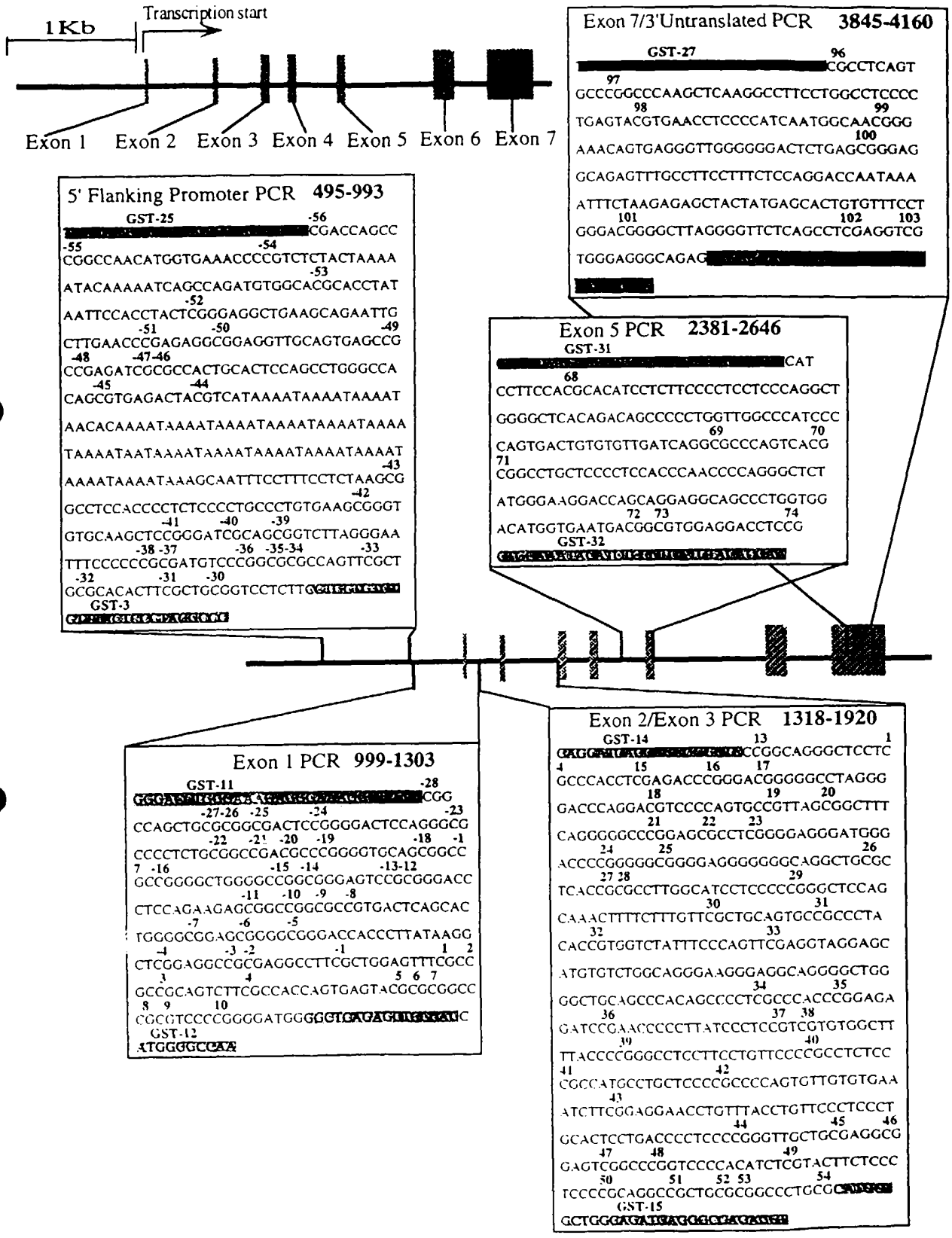

1Kb
Transcription start
Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
Exon 6
Exon 7
Exon 7/3'Untranslated PCR 3845-4160
GST-27
5' Flanking Promoter PCR 495-993
GST-25
GST-3
Exon 5 PCR 2381-2646
GST-31
GST-32
Exon 1 PCR 999-1303
GST-11
GST-12
Exon 2/Exon 3 PCR 1318-1920
GST-14
GST-15

Figure 2A Methylation Status of Individual Sites in the GST-Pi Gene

site		LN	Du	PC3	PC3 M	PC3 MM		2AN	BN CN	2AC	BC	CC	DC	XC	WC	Pr
										4+4	3+3	2+3	2+2	3+4	3+3	
-28		++++	-	+	++	++		-	-	-	++	+	++	++++	++	-
-27		++++	-	+	++	++		-	-	-	+	+	++	++++	+	-
-26		++++	-	+	++	++		-	-	-	+	+	++	++++	++	-
-25		++++	-	+	++	++		-	-	-	+	++	++	++++	++	-
-24		++++	-	+	++	++		-	-	-	+	++	++	++++	+++	-
-23		++++	-	+	+	+		-	-	-	++	++	++	-	+++	-
-22		++++	-	++	+	+		-	-	-	++	+	++	-	+++	-
-21		++++	-	+	-	-		-	-	-	++	+	+++	++++	++	-
-20		++++	-	-	-	-		-	-	-	++	++	++	-	+	-
-19		++++	-	++	-	-		-	-	-	+	++	++	+	+	-
-18		++++	-	++	+	++		-	-	-	++	++	++	++++	++	-
-17		++++	-	+	+	++		-	-	-	++	++	++	++++	+++	-
-16		++++	-	++	+	++		-	-	-	+	++	+++	++++	+++	-
-15		++++	-	+	+	++		-	-	-	++	++	+++	++++	+++	-
-14		++++	-	-	+	+		-	-	-	+	++	++	-	+	-
-13		++++	-	+	+	+++		-	-	-	B	B	B	++	+	-
-12		++++	-	+	++	++		-	-	-	B	B	B	++++	+	-
-11		++++	-	++	++	++		-	-	-	+	++	++	++++	+++	-
-10		++++	-	++	++	++		-	-	-	++	++	+++	++++	++	-
-9		++++	-	++	++	++		-	-	-	++	++	++	++++	+++	-
-8		++++	-	++	+	++		-	-	-	++	+	+++	++++	++	-
-7		++++	-	++	++	++		-	-	-	++	++	+++	++++	+++	-
-6		++++	-	++	++	++		-	-	-	++	++	+++	++++	++	-
-5		++++	-	++	++	+++		-	-	-	++	++	+++	++++	+++	-
-4		++++	-	++	++	+++		-	-	-	++	++	+++	++++	+++	-
-3		++++	-	++	++	++		-	-	-	+++	++	+++	++++	+++	-
-2		++++	-	++	++	B		-	-	-	+++	++	+++	++++	+++	-
-1		++++	-	++	+	B		-	-	-	++	++	++	++++	++	-

Figure 2A (cont'd)

1		++++	-	++	B	B		-	-		++	++	++	++++	+++	-
2		++++	-	++	B	B		-	-		++	++	++	++++	+++	-
3		++++	-	++	B	B		-	-		++	++	++	++++	+++	-
4		++++	-	++	B	B		-	-		++	++	++	++++	+++	-
5		++++	-	+++	B	B		-	-		B	++	+++	++++	++++	-
6		++++	-	+++	B	B		-	-		B	+++	+++	++++	++++	-
7		++++	-	+++	B	B		-	-		B	+++	+++	++++	++++	-
8		++++	-	++	B	B		-	-		B	++	++	++++	+++	-
9		++++	-	++	B	B		-	-		B	++	++	++++	+++	-
10		++++	-	++	B	B		-	-		B	++	++	++++	+++	-
13		++++		+	+++	+++			-		+++	++++	-			
14		++++		+	+++	+++			-		+++	++++	-			
15		++++		+	+++	+++			-		++	++++	-			
16		++++		+	+++	+++			-		++	+++	-			
17		++++		+	++	++			-		++	++++	-			
18		++++		+	++	-			-		+++	++++	-			
19		++++		+	+	++			-		++	+++	-			
20		++++		+	+	+			-		++	++++	-			
21		++++		B	+	+			-		+++	+++	-			
22		++++		B	+	-			-		++	++	-			
23		++++		B	+	-			-		+++	++++	-			
24		+++		B	-	-			-		++	-	-			
25		-		B	+	-			-		++	-	-			
26		++++		B	++	B			-		+++	++++	-			
27		++++		B	+++	B			-		+++	++++	-			
28		++++		B	+++	B			-		+++	++++	-			
29		++++		B	+++	B			-		B	++++	-			
30		++++		B	B	B			-		+++	B	-			
31		++++		B	B	B			-		+++	B	-			
32		B		B	B	B			-		+++	B	-			
33		++++		B	+++	B			-		+++	+++	-			

Figure 2A (cont'd)

68		++++			++++	++++			++++		++++	++++	++++			
69		++++			++++	++++			++++		++++	++++	++++			
70		++++			++++	++++			++++		++++	++++	++++			
71		++++			++++	++++			++++		++++	++++	++++			
72		++++			++++	++++			++++		++++	++++	++++			
73		++++			++++	++++			++++		++++	++++	++++			
74		++++			++++	++++			++++		++++	++++	++++			
96		++++	-	++++	++++	++++			++++		++++	++++	++++			
97		++++	-	++++	++++	++++			++++		++++	++++	++++			
98		++++	-	++++	++++	++++			++++		++++	++++	++++			
99		++++	-	++++	+++	++++			++++		++++	++++	++++			
100		++++	++++	++	++++	++++			++++		+++	+++	++++			
101		++++	++++	++++	++++	++++			++++		++++	++++	++++			
102		++	-	+++	++++	++++			++++		++++	++++	++++			
103		++++	-	++++	++++	++++			++++		++++	+++	+++			

The methylation status at individual CpG sites is shown for the LNCaP (LN), DU145 (DU), PC3, PC3-M and and PC3-MM, for DNA isolated from normal tissue samples from prostate cancer patients (2AN, BN and CN), for prostate tumour tissue (BC, CC, DC, XC, WC and 2AC) and for normal prostate tissue (Pr) from a person without prostate cancer. For prostate tumours the Gleason grade of the tumour is shown on the second row.

The level of methylation at individual sites is indicated:

- \- no methylation
- \+ 1 to 25% methylation
- ++ 26 to 50% methylation
- +++ 51 to 75% methylation
- ++++ 76 to 100% methylation

A blank box indicates that the site was not assayed. A "B" indicates that the status of the site could not be determined, eg. because of a sequence blockage or it was beyond the range of the sequencing run..

Figure 2B Methylation Status of Individual Sites in the GST-Pi Gene

Site	Normal Prostate (15)	PC-3 (10)	LNCaP (7)	BC (9)	CC (4)	DC (10)	Blood (13)	Brain (6)	Spleen (6)	Liver (6)	Smooth muscle (6)	Lung (5)	Bone marrow (5)	Pancreas (6)	Heart (3)
-56	+++	+++	++++	+++	+++	+++	++++	++++	++++	++++	++++	+++	++++	++	++++
-55	++++	+	++++	++	+	+++	++++	++++	+++	++	+++	++	++++	+++	++++
-54	++	-	++++	++	+	++	++++	+++	++	++	++++	++	++++	+++	++++
-53	+++	-	+++	++++	+	+	++++	++++	++++	++++	++++	++++	++++	++	++++
-52	++	+	++++	+	-	++	++++	+	++	+	-	+	++	++	++++
-51	+	-	++++	++	++	+++	++++	++	++	-	+++	++++	+++	++	+++
-50	++	++	++++	++	++	++++	++++	++++	+++	++	+	++++	+++	++	++++
-49	++	+	++++	+	B	+++	++++	++	++	+	-	+++	++++	++	-
-48	+	+++	++++	++	++	++++	+	-	-	-	-	-	-	-	-
-47	+	++	++++	+++	++	+++	+++	++	+++	+	-	+++	++++	+	+
-46	++	++	++++	+++	++	++++	++++	++	++	+	-	+++	+	++	+
-45	++	-	++++	+	++	++++	++++	++	++	+	-	+++	++	+	-
-44	++	-	++++	-	++	+++	++++	+	++	+	-	-	+	++	+++
-43	-	-	++++	+++	++++	++++	-	+	-	-	-	+	-	-	-
-42	-	-	++++	++++	++++	+++	-	-	-	-	-	-	-	-	-
-41	-	-	++++	+++	++++	++++	-	-	-	-	-	-	-	-	-
-40	-	-	++++	++++	++++	++++	-	-	-	-	-	-	-	-	-
-39	-	-	++	++++	++++	++++	-	-	-	-	-	-	-	-	-
-38	-	-	+++	++++	++++	++++	-	-	-	-	-	-	-	-	-
-37	-	-	++++	+++	++++	++++	-	-	-	-	-	-	-	-	-
-36	-	-	++++	++	+	++++	-	-	-	-	-	-	-	-	-
-35	-	+	+++	++++	++++	++++	-	-	-	-	-	-	-	-	-
-34	-	-	++++	++++	++++	++++	-	-	-	-	-	-	-	-	-
-33	-	p	++++	(++++)p	++++	(++++)p	-	-	-	-	-	-	-	-	-
-32	-	-	B	+++	++++	++	-	-	-	-	-	-	-	-	-
-31	-	-	B	++++	++++	++++	-	-	-	-	-	-	+	-	-
-30	-	-	B	+++	++++	++++	-	-	-	-	-	-	-	-	-
-28	-						-	-	-	-	-	-	-	-	-
-27	-						-	-	-	-	-	-	-	-	-
-26	-						-	-	-	-	-	-	-	-	-

Figure 2B (cont'd)

-25	-						-	-	-	-	-	-	-	-	-
-24	-						-	-	-	-	-	-	-	-	-
-23	-						-	-	-	-	-	-	-	-	-
-22	-						-	-	-	-	-	-	-	-	-
-21	-						-	-	-	-	-	-	-	-	-
-20	-						-	-	-	-	-	-	-	-	-
-19	-						-	-	-	-	-	-	-	-	-
-18	-						-	-	-	-	-	-	-	-	-
-17	-						-	-	-	-	-	-	-	-	-
-16	-						-	-	-	-	-	-	-	-	-
-15	-						-	-	-	-	-	-	-	-	-
-14	-						-	-	-	-	-	-	-	-	-
-13	-						-	-	-	-	-	-	-	-	-
-12	-						-	-	-	-	-	-	-	-	-
-11	-						-	-	-	-	-	-	-	-	-
-10	-						-	-	-	-	-	-	-	-	-
-9	-						-	-	-	-	-	-	-	-	-
-8	-						-	-	-	-	-	-	-	-	-
-7	-						-	-	-	+	-	-	-	-	-
-6	-						-	-	-	-	-	-	-	-	-
-5	-						-	-	-	+	-	-	-	-	-
-4	-						-	-	-	+	-	-	-	-	-
-3	-						-	-	-	+	-	-	-	-	-
-2	-						-	-	-	+	-	-	-	-	-
-1	-						-	-	-	+	-	-	-	-	-
1	-						-	-	-	+	-	-	-	-	-
2	-						-	-	-	++	-	-	-	-	-
3	-						-	-	-	++	-	-	-	-	-
4	-						-	-	-	++	-	-	-	-	-
5	-						-	-	-	+	-	-	-	-	-
6	-						-	-	-	+	-	-	-	-	-
7	-						-	-	-	+	-	-	-	-	-
8	-						-	-	-	-	-	-	-	-	-
9	-						-	-	-	-	-	-	-	-	-
10															

Figure 2B (cont'd)

The methylation status at individual CpG sites between -43 and +10 is shown for DNA isolated from normal prostate tissue (from a person without prostate cancer), from three prostate cancer samples (BC, CC and DC) and for a number of other normal tissues.

The level of methylation at individual sites is indicated:

-	no methylation
+	1 to 25% methylation
++	26 to 50% methylation
+++	51 to 75% methylation
++++	76 to 100% methylation

A blank box indicates that the site was not assayed. A "B" indicates that the status of the site could not be determined, eg. because of a sequence blockage or it was beyond the range of the sequencing run.. Patients B and D were polymorphic at CpG site -33 and the level of methylation indicated in the brackets reflects methylation of the allele which contains the CpG.

For sites -28 to +10 the level of methylation was determined by direct sequence analysis of the population of PCR molecules (17). For the upstream CpG sites, -56 to -30, PCR products were cloned and a number of individual clones sequenced (number indicated in brackets below the sample name). For normal tissues the level of methylation at each site was determined as the fraction of all clones containing a C at that position. For the cancer samples BC, CC and DC the level of methylation shown is that among the clones which showed DNA methyaltion in the region from CpG site -43 to -30 (about half of the clones in each case)

Figure 3 **Region of Differential Methylation in Prostate Cancer**

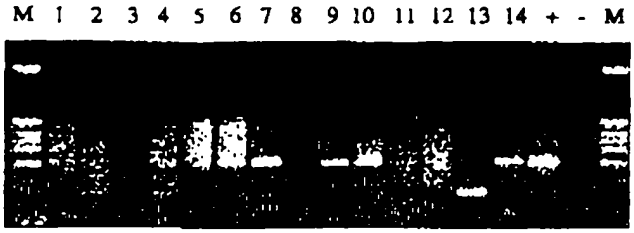
```
                                                  -43                                                  -42
ATAAAATAAA ATAAAATAAA ATAAAGCAAT TTCCTTTCCT CTAAGCGGCC TCCACCCCTC TCCCCTGCCC TGTGAAGCGG -355
ATAAAATAAA ATAAAATAAA ATAAAGTAAT TTTTTTTTTT TTAAGTGGTT TTTATTTTTT TTTTTTGTTT TGTGAAGTGG  B-U
ATAAAATAAA ATAAAATAAA ATAAAGTAAT TTTTTTTTTT TTAAGCGGTT TTTATTTTTT TTTTTTGTTT TGTGAAGCGG  B-M


                                                                                GTA-GC (p)
            -41   -40   -39                          -38-37     -36-35-34     -33   -32
GTGTGCAAGC TCCGGGATCG CAGCGGTCTT AGGGAATTTC CCCCCGCGAT GTCCCGGCGC GCCAGTTCGC TGCGCACACT -275
GTGTGTAAGT TTTGGGATTG TAGTGGTTTT AGGGAATTTT TTTTTGTGAT GTTTTGGTGT GTTAGTTTGT TGTGTATATT  B-U
GTGTGTAAGT TTCGGGATCG TAGCGGTTTT AGGGAATTTT TTTTCGCGAT GTTTCGGCGC GTTAGTTCGT TGCGTATATT  B-M

                 CGPS-5  YGGTTTT AGGGAATTTT TTTTCGC>CGPS-6 YGGYGY GTTAGTTYGT TGYGTATATT
                         CGPS-11  GGGAATTTT TTTTCGCGAT GTTTYGGCGC>


-31  -30                                          -29                                                  -28
TCGCTGCGGT CCTCTTCCTG CTGTCTGTTT ACTCCCTAGG CCCCGCTGGG GACCTGGGAA AGAGGGAAAG GCTTCCCCGG -195
TTGTTGTGGT TTTTTTTTTG TTGTTTGTTT ATTTTTTAGG TTTTGTTGGG GATTTGGGAA AGAGGGAAAG GTTTTTTTGG B-U
TCGTTGCGGT TTTTTTTTTG TTGTTTGTTT ATTTTTTAGG TTTCGTTGGG GATTTGGGAA AGAGGGAAAG GTTTTTTCGG B-M

TC>


      -27-26-25   -24             -23         -22 -21 -20 -19         -18 -17 -16
CCAGCTGCGC GGCGACTCCG GGGACTCCAG GGCGCCCCTC TGCGGCCGAC GCCCGGGGTG CAGCGGCCGC CGGGGCTGGG -115
TTAGTTGTGT GGTGATTTTG GGGATTAAAG GGTGTTTTTT TGTGGTTGAT GTTTGGGGTG TAGTGGTTGT TGGGGTTGGG  B-U
TTAGTTGCGC GGCGATTTCG GGGATTAAAG GGCGTTTTTT TGCGGTCGAC GTTCGGGGTG TAGCGGTCGT CGGGGTTGGG  B-M

      <GCG CCRCTAAARC CCCTAAAATC CCRC CGPS-7    < GCTG CAARCCCCAC ATCRCCARCA RCCCCA CGPS-8
        <G CCGCTAAAGC CCCTAAAATC CCRCAAAA CGPS-12                    <GCCARCA GCCCCAACCC
```

FIGURE 3 (Continued)

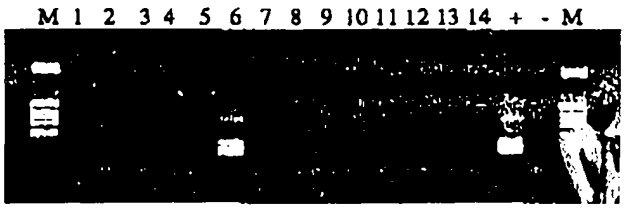
```
 -15 -14     -13-12                  -11 -10  -9 -8                  -7    -6    -5
GCCGGCGGGA GTCCGCGGGA CCCTCCAGAA GAGCGGCCGG CGCCGTGACT CAGCACTGGG GCGGAGCGGG GCGGGACCAC -35
GTTGGTGGGA GTTTGTGGGA TTTTTTAGAA GAGTGGTTGG TGTTGTGATT TAGTATTGGG GTGGAGTGGG GTGGGATTAT B-U
GTCGGCGGGA GTTCGCGGGA TTTTTTAGAA GAGCGGTCGG CGTCGTGATT TAGTATTGGG GCGGAGCGGG GCGGGATTAT B-M

 <GCCGCCCT CAARCRCCCT AAAAAATCTT CTC  CGPS-9 (10)
CAGCCRCCCT CAA  CGPS-13


              -4    -3   -2        -1  >       1  2  3         4                  5 6 7
CCTTATAAGG CTCGGAGGCT GCGAGGCCTT CGCTGGAGTT TCGCCGCCGC AGTCTTCGCC ACCAGTGAGT ACGCGCGGCC +46
TTTTATAAGG TTTGGAGGTT GTGAGGTTTT TGTTGGAGTT TTGTTGTTGT AGTTTTTGTT ATTAGTGAGT ATGTGTGGTT B-U
TTTTATAAGG TTCGGAGGTC GCGAGGTTTT CGTTGGAGTT TCGTCGTCGT AGTTTTCGTT ATTAGTGAGT ACGCGCGGTT B-M

            CGPS-1   C GCGAGGTTTT CGTTGGAGTT TCGTCGTC> CGPS-2 CGTT ATTAGTGAGT ACGCGCGGTT


8 9     10
CGCGTCCCCG GGGATGGGGC TCAGAGCTCC CAGCATGGGG  CCAA +90
TGTGTTTTTG GGGATGGGGT TTAGAGTTTT TAGTATGGGG  TTAA B-U
CGCGTTTTCG GGGATGGGGT TTAGAGTTTT TAGTATGGGG  TTAA B-M

C>
```

FIG 4a

A

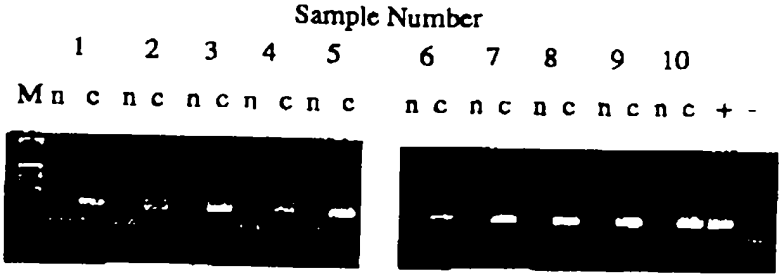


B



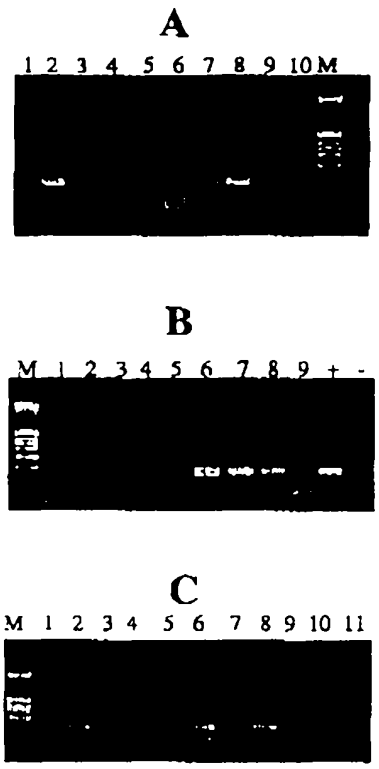
1. Brain, 2. Lung, 3. Skeletal muscle, 4. Spleen, 5. Pancreas, 6. "Normal" Prostate Aged 85y.o., 7. "Normal" Prostate Aged 62y.o., 8. Heart, 9. Bone Marrow, 10. Blood-1, 11. Blood-2, 12, Blood-3, 13.Liver-1, 14. Liver-2.

FIG 4b



Sample	Tissue	Gleason	% Methylation Non CG rich PCR
1	Normal	N/A	-
	Cancer	3+3	++++
2	Normal	N/A	-
	Cancer	3+5	++
3	Normal	N/A	-
	Cancer	3+3	++
4	Normal	N/A	-
	Cancer	3+5	-
5	Normal	N/A	-
	Cancer	2+2	++
6	Normal	N/A	-
	Cancer	3+3	-
7	Normal	N/A	-
	Cancer	2+3	++
8	Normal	N/A	-
	Cancer	3+3	++
9	Normal	N/A	-
	Cancer	2+3	++++
10	Normal	N/A	-
	Cancer	?	++

FIG 4c



A Blood samples obtained from prostate cancer patients during radical prostacectomy

B 1. Normal prostate-1, 2. Normal prostate-2, 3. Normal prostate-3, 4. Normal prostate-4, 5. Normal prostate-5, 6. HPV transformed prostate cell line, 7. Blood from prostate patient PA (PSA=1000), 8. Blood from prostate patient PB (PSA=56), 9. Blood from prostate patient PC (PSA=18).

C 1. LNCaP cell line, 2. Du145 cell line, 3. PC-3 cell line, 4. PC-3M cell line, 5. PC-3MM cell line, 6. Hela cell line, 7. Leukemic DNA, 8. HepG2 cell line, 9. Human liver DNA, 10. White blood cells 11. MRC-5 cell line